《AI 营销智能体交互过程与效果有效性评估规范》标准编制说明

中国信息通信研究院标准起草组

2026 年 11 月 2 日

1、 标准范围。

本文件规定了AI营销智能体交互过程能力与效果有效性的能力要求，以及评估方法、评估流程和评估结果等级划分等要求。

本文件适用于AI营销智能体研发方开展产品自评估、使用方开展选型评估，以及第三方测评机构开展第三方评估活动。

2、 工作简况。

本项目计划名称为“AI智能体营销交互过程与效果有效性评估规范”，文件名称为“AI营销智能体交互过程与效果有效性评估规范”。由中国互联网协会归口。由中国信息通信研究院和北京执牛耳传媒科技有限公司牵头研制。

本文件起草单位包括中国信息通信研究院、北京执牛耳传媒科技有限公司、景智光年（上海）科技服务有限公司、旺脉集团、北京特比昂科技有限公司、上海星谷信息科技有限公司。

本文件主要起草人：×××、×××、×××、×××、×××、×××。

本文件于2026年11月在中国互联网协会完成立项审核，立项申报材料中的项目计划起止时间自2026年3月起。

2026年3月，起草组成立并启动标准研制工作，梳理国内外AI营销智能体产品形态、技术路线与评估实践，确定“能力要求+评估”的标准框架，形成标准草案初稿。

2026年5月，中国信息通信研究院联合举办AI智能体营销交互过程与效果有效性评估规范研讨会，就标准定位、能力维度划分与评估方法听取产业界意见，据此完善标准草案。

2026年11月2日，中国互联网协会组织立项审核专家对标准进行审核，沈昌祥、邹潇湘、吕军三位专家共提出10条意见，主要涉及术语与缩略语精简、能力

要求表述统一、评估框架图示、功能点属性与等级划分、引言与范围表述等方面。起草组对意见逐条研究处理，采纳9条、未采纳1条，处理情况见《专家意见汇总表（含修改说明）》，并形成本标准文本及本编制说明。

3、 标准编制原则和确定标准主要内容的依据：

编制原则：

- 规范性要求，按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草，条款表述、层次划分与要素编排统一规范，并充分借鉴国内相关标准规范，力求标准的可执行性和规范性更强；
- 实用性要求，建立在对AI营销智能体产业实践与典型营销场景充分调研的基础之上，能力要求与功能点设置直接对应研发、选型与测评的实际需要，具有充分的技术先进性和实用性；
- 可行性要求，针对AI营销智能体产品形态差异大、交互行为难以穷举的特点，以功能点验证替代传统测试用例，使评估结论可观察、可验证、可复现；
- 协调性要求，与GB 45438—2025等已发布标准以及生成式人工智能相关管理规定保持协调，不重复规定、不与强制性标准冲突；
- 有效性要求，考虑到技术与业务形态的持续演进，采用“能力要求+评估”的开放式框架，功能点清单可按行业与场景增补调整，满足未来标准发展与扩充需求。

确定标准主要内容的依据：

- 关于范围（第1章）：界定本文件规范的对象与适用主体。考虑到其他类型智能体的营销相关评估在能力维度与指标口径上与本文件存在差异，

专家审核后删去“其他类型智能体的营销相关评估活动可参照使用”的表述，避免适用范围泛化导致评估结论不可比。

- 关于术语和定义（第3章）：仅界定本文件反复使用且易产生歧义的概念，共6条。其中“效果有效性”界定为“在营销效果、交互效率、用户体验与信任、商业价值等方面所达到的可量化成效水平”，明确本文件评价的是量化水平而非二值判定，据此未采纳将其更名为“效果效率”的意见。
- 关于缩略语（第4章）：仅保留正文实际使用的AI、CRM两项，删除LLM、CSAT、NPS、ROI等未在正文使用的缩略语，避免正文与缩略语表不对应。
- 关于能力要求（第5章）：按“交互过程能力+效果有效性能力”双维度设置。5.2从意图理解与识别、对话交互、高质量营销内容生成、任务执行与工具调用、人机协同与转接、合规与安全六个方面提出要求，覆盖营销交互全链条；5.3从营销效果、交互效率、用户体验与信任、商业价值四个方面提出指标要求，并单列5.3.5效果指标披露与可验证性要求，解决效果数据由谁提供、口径是否可核的问题，为评估结论的可追溯性提供支撑。
- 关于评估（第6章）：明确评估目的与原则、评估框架、评估方式、功能点验证、评估流程与等级划分。评估以功能点验证为主，将第5章能力要求细分为44个可验证功能点（交互过程能力27个、效果有效性能力17个），并按必需、增强两类属性区分；验证方式包括材料审查、系统检测、专家测评、用户调研和场景测试；结果划分为完备级、优秀级、卓越级三个等级，分别对应“全部必需功能点通过”“全部必需功能点及不少于50%的增强功能点通过”“全部必需功能点及全部增强功能点通过且关键能力

表现突出”。为提高等级区分度，必需与增强功能点的数量比例宜控制在7:3左右，现行为30:14。

- 关于附录A：以资料性附录给出功能点验证记录与评估报告的参考模板，统一证据留存与记录格式，便于评估结果互认。

4、 主要试验（或验证）的分析、综述报告。

本文件为评估类标准，不涉及产品设计、生产工艺类试验，主要验证工作为功能点的可操作性与等级划分的可区分性。起草组基于对参与单位AI营销智能体产品的调研与试用，选取客户触达、需求识别、产品推荐、商机培育、售后跟进等典型营销场景，对功能点逐项开展内部试评，重点验证三方面内容：一是功能点是否可观察、可判定，避免出现“是否具备某某能力”这类无法给出明确结论的表述；二是证据要求是否可落地，明确材料审查、系统检测、专家测评、用户调研、场景测试五类方式与证据形态的对应关系；三是等级划分是否具备区分度，试评发现原清单中必需功能点占比过高（44项中40项为必需），优秀级与卓越级之间可供区分的能力项过少，据此调整了必需与增强功能点的属性配比，并将优秀级阈值确定为增强功能点总数的50%。

上述验证为起草组内部验证，用于确认标准条款的可操作性，不作为对外评估结论。标准发布后，将通过试点评估进一步积累数据，必要时对功能点清单与等级阈值作出修订。

5、 标准在起草过程中遇到的问题及解决办法：重大分歧意见的处理经过和依据：有无重要技术问题需要说明。

起草过程中遇到的主要问题及解决办法如下：

- AI营销智能体产品形态、业务场景与任务类型差异较大，传统测试用例方式难以覆盖千差万别的交互行为。解决办法：以功能点验证为主，将

能力要求细分为可观察、可验证的功能点，并以功能点验证结果作为评估与等级判定的主要依据。

- “交互过程”与“效果有效性”两个维度的边界容易混淆。解决办法：明确5.2规定交互过程中的能力表现、5.3规定交互所产生的可量化结果与价值，二者并列构成评估框架，条款不重复设置。
- 等级划分区分度不足。解决办法：将体验提升类、展示类与条件性要求调整为增强功能点，同步将对应条款的限定词由“应”改为“宜”，使表属性与条款表述保持一致；并补充必需与增强功能点配比宜为7:3的要求。
- 效果数据来源与统计口径不统一，评估结论难以核实。解决办法：设置5.3.5效果指标披露与可验证性要求，对指标口径、数据来源、统计周期、对比基准与支撑数据的可追溯性提出要求。

关于重大分歧意见：专家提出将标准名称及正文中的“效果有效性”调整为“效果效率”。经研究未予采纳，主要考虑：本文件评价的是营销效果所达到的量化水平，属程度性判断，与术语3.3的定义一致；“效率”通常指向投入产出比与响应速度，已由5.3.2交互效率指标覆盖，整体更名将与指标维度产生语义重叠；该名称亦已在立项申报材料中沿用，更名涉及范围、术语、附录等多处联动修改。起草组已就该意见与专家沟通说明。

除上述一条外，无其他重大分歧意见。标准起草过程中未发现需要专门说明的重要技术问题。

6、 与国外标准的关系：包括：采用国际标准和国外先进标准的程度，与国外标准主要技术内容的差异（可引用标准前言的内容）：

本文件未采用国际标准或国外先进标准，无对应的国际标准。

目前ISO/IEC JTC 1/SC 42已发布人工智能管理体系、风险管理等相关标准，有关机构也发布了人工智能风险管理框架，但均面向人工智能通用治理与风险，尚未形成面向AI营销智能体交互过程与效果有效性的专用评估标准。本文件在评估思路上与国际通行的“要求—证据—分级”模式一致，在能力维度设置上结合我国营销业务实际，突出营销内容合规、生成内容标识与个人信息保护等要求。

7、 修订标准时，说明与标准前一版本的重大技术变化，并列出所涉及的新、旧版本的有关章条（可引用标准前言的内容）：废止/代替现行有关标准的建议：

本文件为首次制定，无前一版本，不涉及新旧版本技术内容的对比。

无废止或代替现行标准的建议。

8、 说明标准与其他标准或文件的关系（可引用标准前言的内容），特别是与有关的现行法律、法规和强制性国家标准的关系：

本文件规范性引用了GB 45438—2025《网络安全技术 人工智能生成合成内容标识方法》，营销内容的生成与标识要求与其保持一致。

编制过程中参考了T/CCSA 737-2025（T/CAAAD 005-2025）《人工智能营销客服平台能力要求》等已发布标准，在能力维度划分上与其衔接，不重复规定其已有内容。

本文件与《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》《生成式人工智能服务管理暂行办法》《互联网信息服务深度合成管理规定》《人工智能生成合成内容标识办法》等法律法规、部门规章的要求一致，不存在冲突；与现行强制性国家标准无冲突。

9、 标准作为强制性标准或推荐性标准的建议：

建议本文件作为推荐性团体标准发布实施。

10、 贯彻国家标准的要求和措施建议（包括组织措施、技术措施、过渡办法等内容）：标准发布后，对国内外业界可能产生的影响。

组织措施：建议由中国互联网协会组织标准宣贯培训，面向AI营销智能体研发企业、使用单位和第三方测评机构开展条款解读，培育具备相应能力的评估机构，推动评估队伍建设。

技术措施：配套提供功能点清单与评估报告模板（见附录A）的电子化版本，统一证据留存与记录格式，便于评估结果在机构之间互认。

过渡办法：标准发布后设置6个月过渡期，期间以试点评估方式先行验证，评估机构可结合行业与业务场景确定适用的功能点清单，并在评估报告中披露判定依据；过渡期结束后转为常态实施。

影响：本文件为供需双方提供统一、可量化、可验证的评估标尺，有助于改变“重建设、轻评估”的状况，降低选型与采购的沟通成本，推动AI营销智能体从功能堆叠转向效果可证；对国外业界，可为AI营销领域的评估规则提供中国实践参考，也有利于我国相关产品在跨境营销场景中的能力说明与结果互认。

11、 标准是否涉及知识产权的情况说明；如标准中含有自主知识产权，说明产品研发程度、产业化基础及进程。

本文件不涉及知识产权问题，标准中不包含专利技术。

12、 其他应予说明的事项。

本文件于2026年11月2日完成立项审核专家意见的处理与修改，专家意见共10条，采纳9条、未采纳1条，逐条处理情况见《专家意见汇总表（含修改说明）》。

本文件尚需履行征求意见、送审、报批等程序；文中目次页码、起草单位与起草人信息待正式报送前确认并更新。